SZEMEREDI'S REGULARITY THEOREM; FROM AN ANALYTIC POINT OF VIEW

by

Levent Bilir

Bachelor of Science, in Mathematics, Boğaziçi University, 1990

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Mathematics
Boğaziçi University
2010

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Assoc. Prof. S. Demir for his invaluable guidance and help throughout my studies. This project would never be completed without his constructive comments and suggestions.

I am grateful to Prof. A. Ülger for his moral support and encouragement.

My appreciation is also extended to the members of my thesis committee for allocating their time to evaluate this work.

Finally, in addition to my parents, I need to admit that I was indebted to a number of people for their understanding and patience. Their support during the tough times, made the real difference.

ABSTRACT

SZEMEREDI'S REGULARITY THEOREM; FROM AN ANALYTIC POINT OF VIEW

The aim of this thesis is to study Szemerédi's Regularity Lemma from an analytical point of view. We are going to explain how one can interpret the so-called Szemeredi's Regularity Lemma in the language of analysis. By using this language, we are going to prove two analytical statements which imply several (both analytical and graph theoretical) forms of Szemeredi's Regularity Lemma. For the sake of completeness, we will present a combinatorial proof of the standard form of this lemma in the appendix.

ÖZET

ANALİTİK YAKLAŞIMLA SZEMEREDİ'NİN REGÜLERİTE YARDIMCI TEOREMİ ÜZERİNE İNCELEME

Szemerédi Regülerite Yardımcı Teoremi'nin, analiz dilinde nasıl yorumlanabileceğini açıklayacağız. Daha sonra da, bu dili kullanarak Szemerédi Regülerite Yardımcı Teoremi'nin değişik versiyonlarını ve bunları gerektiren bazı teoremler kanıtlayacağız. Bir bütünsellik sağlaması için, bu yardımcı teoremin kombinatorik kanıtını da ek kısımda sunacağız.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

$\mathcal{A}$	Algebra of sets
(a_{ij})	Adjaceny matrix of the graph G
C, C_i	Exceptional class,the set of left over elements with respect to a partition
$d_G\,(X, Y)$	Density of $X - Y$ edges of G
$d_W, d_W^1, d_{\tilde{W}}^1$	Distance functions
E	Set of edges of a graph
$e_G\,(X, Y)$	The number of $X - Y$ edges of the graph G
$e_P\,(S, T)$	Expected number of edges of G connecting S to T
$G = (V, E)$	The graph with vertex set V and the set of edges E.
$G(V_i, V_j)$	Induced bipartite graph between $V_i$and V_j
G^c	Complement of a graph G
$irreg_G\,(U, W)$	Irregularity between U and W
K_i	Arbitrary nonempty subset of a Hilbert Space
$\mathbb{N}$	The set of natural numbers
$\mathcal{P}$	Partition of a set of vertices V
$\mathcal{P}^{'}$	Partition of a set of vertices V refined from $\mathcal{P}$
$\mathcal{P}_i$	Partition of a set V that is the i-th refinement of the initial partition
$q, q^{'}$	Special functions defined and used in the appendix
$R, R^{'}, R^{''}$	Measurable rectangles in $[0, 1]^2$
$\mathbb{R}$	the set of real numbers
$\mathbb{R}[x]$	Ring of polynomials with real coefficients
$S \times T$	Product of sets S and T
S_a	For $S \subseteq V$, the corresponding element in $\mathcal{A}$
s_r	The number of elements of an equal part in an equitable partition $\mathcal{P}_r$
T_W	The kernel operator corresponding to W
$U, W, W^{'}, W^*$	Step functions

V, V_i	Set of vertices
$\mathcal{W}$	The space of all bounded, symmetric, real valued and measurable functions defined on $[0, 1]$
$\mathcal{W}_0$	Subspace of $\mathcal{W}$ whose elements' images are in $[0, 1]$
W_G	A special step function
$(W_G)_{\mathcal{P}}, W_{\mathcal{P}}$	Step function defined as the average value on the parts of $\mathcal{P}$
$\mathbb{Z}$	The set of integers

α, β	Real Numbers
α_i, γ_i	Real numbers
ε	A positive real number that is less than 1
μ	Lebesgue measure of a set
ϕ	Function
ϕ^{-1}	Inverse function of ϕ
φ	Function
χ_s	Characteristic function defined on S

$<$	Less than
$\leq$	Less than or equal to
$>$	greater than
$\geq$	Greater than or equal to
$=$	Equal to
$\Rightarrow$	Implies
$[0, 1]$	Closed interval from 0 to 1
$\subset$	Subset of
$\supset$	Contains
$\subseteq$	Subset or equal to
$\supseteq$	Contains or equal to
$\subsetneq$	Proper subset of
$\circ$	Composition
$\lceil \ \rceil$	Integer that is greater than or equal to
$\lfloor \ \rfloor$	Integer that is less than or equal to

$\lvert\cdot\rvert$	Absolute value of a number, the number of elements of a set
$\langle\cdot,\cdot\rangle$	Inner product
$\lVert\cdot\rVert_1$	$L_1 - norm$
$\lVert\cdot\rVert_2$	$L_2 - norm$
$\lVert\cdot\rVert_\square$	Cut norm
$\lVert\cdot\rVert$	Depending on the context, norm
inf	Infimum
max	Maximum
RLHS	Regularity lemma in Hilbert Space
SRL	Strong Regularity Lemma
sup	Supremum
wlog	Without loss of generality
WRL	Weak regularity Lemma

1. INTRODUCTION

In 1975, the result that every set of the natural numbers with positive upper density contains arbitrarily long arithmetic progressions was proved. A lemma which is an ingredient in proving this theorem, turned out to be of vital importance in solving a great variety of extremal problems so that a special name is granted to it. Today, it is known as Szemeredi's Regularity Lemma. (Due its importance, some mathematicians are naming it as theorem). This Lemma, roughly says that for a given graph G with vertex set V and the set of edges E, there exists a partition of V into V_1,V_2...,V_k such that for almost every pair (i, j), the induced bipartite graph $G(V_i, V_j)$ behaves like a typical random graph of the same density.

There are some combinatorial proofs, even in some textbooks. Within the last decade, a numerous well-known mathematicians (eg. T. Tao, T. Gowers, H. Furstenberg, B. Green, L. Lovázs) tried to simplify the proof, find other proofs or apply the results of the lemma to the other branches of mathematics (Graph Property Testing, Extremal Graph Theory, Combinatorial Number Theory, etc.).

In his complicated proof, Tao used probability and information theory (Tao, 2006). Gowers' combinatorial proof was very complicated (Gowers, 2007). Even if their proofs were applicable to other forms, they are still unclear how to.

Our goal is to try to understand Szemeredi's Regularity Lemma using the language of functional analysis. It is because we know functional analysis better and it seems that some geometric statements in the language of functional analysis imply several forms of The Regularity Lemma. In fact, a functional analytic proof was given by Lovazs and Szegedy (Lovasz and Szegedy, 2007). However this article is not clear to understand, at least for us. In our opinion, examining this article thoroughly will provide us a strong tool to attack the problem of generalizing this theorem to higher dimensions. For the time being, this is beyond the scope of our study.

Graph theoretical version can be considered well known. In this project, in addition to a standard combinatorial proof, two proofs from the analytical point of view are given. We hope that this kind of understanding could shed some light on higher dimensional cases (version for the hypergraphs), which seems to be mysterious.

2. GRAPH THEORETICAL VERSIONS OF SZEMERÉDI'S REGULARITY LEMMA

Firstly, we need to have some notations and definitions:

Given a graph $G = (V, E)$ and a pair (X, Y) of disjoint non-empty subsets of V, denote the number of $X - Y$ edges of G by $e(X, Y) = e_G(X, Y)$ and we write

$$d(X, Y) = d_G(X, Y) = \frac{e(X, Y)}{|X|\,|Y|}$$

for the density of the $X - Y$ edges of G.

Definition 2.1. *Given a graph $G = (V, E)$ and a pair (X, Y) of disjoint non-empty subsets of V, (X, Y) is said to be an $\varepsilon -$ regular pair if $|d(U, W) - d(X, Y)| < \varepsilon$ whenever $U \subseteq X$ and $W \subseteq Y$ are with the condition that $|U| \geq \varepsilon \cdot |X| > 0$ and $|W| \geq \varepsilon \cdot |Y| > 0$.*

Definition 2.2. *A partition $\{C_0, C_1, ..., C_k\}$ of the vertex set V is said to be an equitable partition with exceptional class C_0 if $|C_1| = |C_2| = ... = |C_k|$.*

Definition 2.3. *An equitable partition $P = (C_i)_{i=0}^{k}$ is said to be an $\varepsilon-$regular partition if the exceptional class C_0 has at most $\varepsilon\cdot |V|$ vertices and with all but at most $\varepsilon \cdot k^2$ pairs, the pairs (C_i, C_j) where $1 \leq i < j \leq k$ are ε-regular.*

Definition 2.4. *Let G be a bipartite graph with bipartition $\{U, W\}$. Irregularity between U, W is given as follows;*

$$irreg_G(U, W) = \max_{X \subseteq U, Y \subseteq W} |e_G(X, Y) - d \cdot |X|\,|Y||\,.$$

Definition 2.5. *For a partition $P = \{V_1, ...V_k\}$ of V, define $d_{ij} = \frac{e_G(V_i, V_j)}{|V_i| \cdot |V_j|}$. For any* *two sets* $S, T \subseteq V(G)$*, expected number of edges of* G *connecting* S *to* T *is*

$$e_P\left(S,T\right) = \sum_{i=1}^{k}\sum_{j=1}^{k} d_{ij}\cdot \left|V_i \cap S\right| \left|V_j \cap T\right|.$$

Lemma 2.6 (Usual form)**.** *For every* $\varepsilon > 0$ *and* $m > 0$*, there is an integer* $M = M(\varepsilon, m)$ *such that every graph* $G = (V, E)$ *on at least* k *nodes has an equipartition* $\{V_1, V_2, ...V_k\}$ *with* $m \leq k \leq M$ *such that for all but* εk^2 *pairs of indices* $1 \leq i < j \leq k$*, the bipartite graph* $G\left[V_i, V_j\right]$ *is* ε*-regular.*

Lemma 2.7 (Second form)**.** *For every* $\varepsilon > 0$*, there is a* $k(\varepsilon) > 0$ *such that every graph* $G = (V, E)$ *has an equipartition* P *into* $k \leq k(\varepsilon)$ *classes* $V_1, ..., V_k$ *such that*

$$\sum_{1\leq i<j\leq k} irreg_G\left(V_i, V_j\right) \leq \varepsilon |V|^2.$$

For the next version, which is called as third form, we need to introduce two more functions on the set vertices V;

Definition 2.8. *Let* $G = (V, E)$ *be a graph. For* $u, v \in E$ *let*

$$a_G(u,v) = \begin{cases} 1, & \text{if } uv \in E \\ 0, & \text{otherwise} \end{cases}$$

Definition 2.9. *For a partition* $P = \{V_1, ..., V_k\}$ *of* V *and* $u, v \in V$*, let* $a_P(u, v) = d_G(V_i, V_j)$ *where* $u \in V_i$ *and* $v \in V_j$*.*

Lemma 2.10 (Third form)**.** *For every* $\varepsilon > 0$*, there is a* $k\left(\varepsilon\right) > 0$ *such that every graph* $G = (V, E)$ *has an equipartition* P *into* $k\left(\varepsilon\right)$ *classes* $V_1, ..., V_k$ *such that*

$$\left| \sum_{uv\in E(H)} \left(a_G(u,v) - a_P(u,v)\right) \right| \leq \varepsilon$$

for every graph H *on* V *that is the union of at most* k^2 *complete bipartite graphs.*

Lemma 2.11 (Weak Regularity Lemma)**.** *For every* $\varepsilon > 0$ *and every graph* $G = (V, E)$*,* V *has a partition* P *into* $k \leq 2^{2/\varepsilon^2}$ *classes* $V_1, .., V_k$ *such that for all* $S, T \subseteq V$ *we have* $|e_G(S, T) - e_p(S, T)| \leq \varepsilon |V|^2$.

Claim: $i)$**Lemma 2.6**$(Usual form) \Rightarrow$ **Lemma 2.7**$(SecondForm)$

$ii)$**Lemma 2.10**$(ThirdForm) \Rightarrow$ **Lemma2.7**$(SecondForm)$

Fact: Let A, B be two disjoint subsets of vertices of a graph $G = (V, E)$ and $X \subset A$, and $Y \subset B$. We have the following;

$$|d(A, B) - d(X, Y)| \leq 1 - \frac{|X|}{|A|} + 1 - \frac{|Y|}{|B|}.$$

Proof of the fact: Let $\dfrac{|X|}{|A|} = \alpha > 0$ and $\dfrac{|Y|}{|B|} = \beta > 0$. Then

$$\begin{aligned} e(A, B) - e(X, Y) &\leq |A|\,(|B| - |Y|) + |B|\,(|A| - |X|) \\ &= |A|\,|B|\,(1 - \alpha + 1 - \beta). \end{aligned}$$

By using this, we have

$$\begin{aligned} d(A, B) - d(X, Y) &= \frac{e(A, B)}{|A|\,|B|} - \frac{e(X, Y)}{|X|\,|Y|} \\ &\leq \frac{e(A, B) - e(X, Y)}{|A|\,|B|} \\ &\leq \frac{(1 - \alpha + 1 - \beta)\,|A|\,|B|}{|A|\,|B|} \\ &= 1 - \alpha + 1 - \beta. \end{aligned}$$

So,

$$d(A,B) - d(X,Y) \leq 1 - \frac{|X|}{|A|} + 1 - \frac{|Y|}{|B|}.$$

Now, let G^c denote the complement of the graph $G = (V, E)$. Then

$$\begin{aligned} d(X,Y) - d(A,B) &= 1 - d_{G^c}(X,Y) - (1 - d_{G^c}(A,B)) \\ &= d_{G^c}(A,B) - d_{G^c}(X,Y) \\ &\leq 1 - \frac{|X|}{|A|} + 1 - \frac{|Y|}{|B|}. \end{aligned}$$

Therefore,

$$|d(A,B) - d(X,Y)| \leq 1 - \frac{|X|}{|A|} + 1 - \frac{|Y|}{|B|}.$$

Proof. (Usual Form ⇒ Second form) Using the fact above, for a pair (V_i, V_j) in an equipartition with $X \subseteq V_i$ and $Y \subseteq V_j$, we have

$$|d(X,Y) - d(V_i,V_j)| \leq 1 - \frac{|X|}{|V_i|} + 1 - \frac{|Y|}{|V_j|}.$$

So

$$\begin{aligned} |e(X,Y) - d(V_i,V_j)\,|X|\,|Y|| &\leq \left(1 - \frac{|X|}{|V_i|} + 1 - \frac{|Y|}{|V_j|}\right)|X|\,|Y| \\ &= \left(1 - \frac{|X|}{|V_i|}\right)\frac{|X|}{|V_i|}\,|V_i|\,|Y| + \left(1 - \frac{|Y|}{|V_j|}\right)\frac{|Y|}{|V_j|}\,|V_j|\,|X|. \end{aligned}$$

Put

$$\alpha = \frac{|X|}{|V_i|}.$$

It is clear that $(1-\alpha)\alpha$ attains its maximum when $\alpha = \frac{1}{2}$. Since V_i, V_j are the parts of the same equipartition, $|V_i| = |V_j|$. So,

$$\begin{aligned} |e(X,Y) - d(V_i,V_j)\,|X|\,|Y|| &\leq \frac{1}{4}|V_i|\,|Y| + \frac{1}{4}|V_j|\,|X| \\ &\leq \frac{1}{4}|V_i|\,(|X|+|Y|) \\ &\leq \frac{1}{4}|V_i|\,(|V_i|+|V_j|) \\ &\leq \frac{1}{2}|V_i|\,|V_j|. \end{aligned}$$

Now, let us analyze the irregularity of the pairs of an equipartition. For an $\varepsilon - regular$ pair (V_i, V_j), we have $|e(X,Y) - d(V_i,V_j)\,|X|\,|Y|| \leq \varepsilon\,|X|\,|Y| \leq \varepsilon\,|V_i|\,|V_j|$. If (V_i, V_j) is not $\varepsilon - regular$, then by the inequality we found above, we have

$$|e(X,Y) - d(V_i,V_j)\,|X|\,|Y|| \leq \frac{1}{2}|V_i|\,|V_j|.$$

The number of all the distinct pairs for the partition is $\frac{k(k-1)}{2}$ and (by our assumption) the total number of not $\varepsilon - regular$ pairs is at most εk^2, we have

$$\begin{aligned} \sum_{1\leq i<j\leq k} \mathrm{irreg}_G(V_i,V_j) &= \sum_{\substack{1\leq i<j\leq k \\ (V_i,V_j)\ \varepsilon\text{-regular}}} \mathrm{irreg}_G(V_i,V_j) + \sum_{\substack{1\leq i<j\leq k \\ (V_i,V_j)\ \text{not}\ \varepsilon\text{-regular}}} \mathrm{irreg}_G(V_i,V_j) \\ &\leq \left(\frac{k^2-k}{2}\right)\varepsilon\,|V_i|\,|V_j| + k^2\varepsilon\frac{1}{2}|V_i|\,|V_j| \\ &\leq \varepsilon k^2\,|V_i|\,|V_j| \\ &\leq \varepsilon\,|V|^2. \end{aligned}$$

□

Proof. (Third form ⇒ Second Form) Let us assume that The Third Form is true. For a pair (V_i, V_j) in the equipartition, let $X \subseteq V_i$ and $Y \subseteq V_j$ attain the maximum in the definition of $irreg_G(V_i, V_j)$. Also, let H_{ij} be a complete bipartite graph between X and

Y. Then, we have

$$\begin{aligned}\left|\sum_{uv\in E(H_{ij})}(a_G(u,v)-a_P(u,v))\right| &= \left|e_G\left(X,Y\right)-d\cdot\left|X\right|\left|Y\right|\right| \\ &= \max_{X\subseteq V_i, Y\subseteq V_j}\left|e_G\left(X,Y\right)-d_G(V_i,V_j)\cdot\left|X\right|\left|Y\right|\right| \\ &= irreg_G(V_i,V_j).\end{aligned}$$

By our assumption,

$$\left|\sum_{uv\in E(H_{ij})}(a_G(u,v)-a_P(u,v))\right|\leq\varepsilon.$$

We know that there are

$$\frac{k(k-1)}{2}$$

distint pairs (V_i, V_j) in the equipartition and for each pair we have a complete bipartite graph H_{ij} constructed above which implies the corresponding pair's irregularity is less than or equal to ε. So,

$$\sum_{1\leq i<j\leq k} irreg_G(V_i,V_j)\leq\frac{k(k-1)}{2}\varepsilon\leq\frac{k^2}{2}\varepsilon\leq\varepsilon\left|V\right|^2.$$

$\square$

3. ANALYTIC LANGUAGE

We want to translate the problem in functional analysis. Let $\mathcal{W}$ be the space of all bounded symmetric measurable functions $W : [0,1]^2 \longrightarrow \mathbb{R}$. Let $\mathcal{W}_0$ denote the subspace of $\mathcal{W}$ whose values are in $[0,1]$. We call a function U a step function with at most n steps if there is a partition $\{S_1, S_2, \ldots, S_n\}$ of $[0,1]$ such that U is constant on each $S_i \times S_j$. From the analytic point of view, we consider graphs as $0-1$ valued step functions in $\mathcal{W}_0$ such that all S_i's have the same measure.

Every $W \in \mathcal{W}$ can be considered as a kernel operator on the Hilbert space $L^2([0,1]^2)$ defined by

$$W(f)(x) := \int_0^1 W(x,y)f(y)dy.$$

We have some standard norms on $\mathcal{W}$: L^1 and L^2-norms. We also have

$$\|W\|_\Box := \sup \left| \int_{S\times T} W(x,y)dxdy \right|$$

where the supremum is taken over all the pair of measurable subsets $S, T \subset [0,1]$.

Note. We are interested in non-oriented and simple graphs. In order build up some intuition, as an example, consider the following figures;

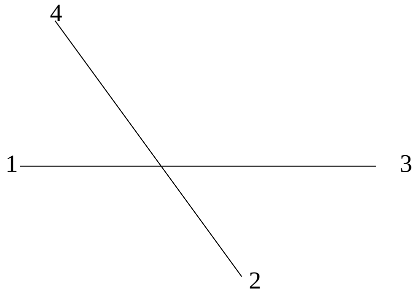

Figure 3.1. Graph $G(V,E)$

$$\begin{array}{c|cccc|} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \\ 3 & 1 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 & 0 \\ \hline \end{array}$$

Figure 3.2. Adjacency matrix of the graph G

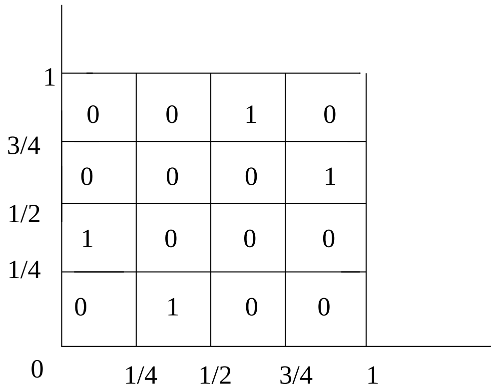

Figure 3.3. $W_G : [0,1] \times [0,1] \to \mathbb{R}$

For the function corresponding to the given graph in Figure 3.3, we divide $[0,1]$ into equal parts with the partition size equal to $|V|$. There is a one to one correspondence between the vertices of the graph and the parts of $[0,1]$. We assign that corresponce with respect to the adjacency matrix. As you see, this function, whose graph appears in Figure 3.3, is defined on parts and it is a characteristic function on each part. We have values 0 or 1, depending on whether or not there exist an edge between the corresponding vertices.

3.1. Norms on $\mathcal{W}$

(i) Since each $F \in \mathcal{W}$ is bounded and $\mu([0,1]^2) = 1 < \infty$, we have $F \in L^1([0,1])$. Hence we may consider $\mathcal{W}$ as a subspace of $L^1([0,1])$. When we write $\|W\|_1$ we mean the norm of W as an element of $L^1([0,1])$.

(ii) Similarly, one has

$$\int_{[0,1]^2} |F|^2 d\mu < \infty.$$

Hence, we may consider $\mathcal{W}$ as a subspace of $L^2([0,1])$ as well.

(iii) Finally, the most important norm for us is the so-called cut norm. It is defined as follows:

$$\|F\|_\Box := \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} F(x,y)dxdy \right|.$$

To show that $\|.\|_\Box$ is indeed a norm, it is obvious that verifying the triangle inequality would be enough;

If $F, G \in \mathcal{W}$ and $\varepsilon > 0$, then (by definition) there exist some measurable sets $S, T \subseteq [0,1]$ such that

$$\|F+G\|_\Box - \varepsilon < \left| \int_{S\times T} (F+G)(x,y)dxdy \right|.$$

Now, it is straightforward to verify that

$$\begin{aligned} \left| \int_{S\times T} (F+G)(x,y)dxdy \right| &= \left| \int_{S\times T} F(x,y)dxdy + \int_{S\times T} G(x,y)dxdy \right| \\ &\leq \left| \int_{S\times T} F(x,y)dxdy \right| + \left| \int_{S\times T} G(x,y)dxdy \right| \\ &\leq \|F\|_\Box + \|G\|_\Box \end{aligned}$$

because

$$\|F\|_{\square} + \|G\|_{\square} = \sup_{S,T\subseteq[0,1]} \left|\int_{S\times T} F(x,y)dxdy\right| + \sup_{S,T\subseteq[0,1]} \left|\int_{S\times T} G(x,y)dxdy\right|.$$

Since $\varepsilon > 0$ was arbitrary, this shows that $\|F+G\|_{\square} \leq \|F\|_{\square} + \|G\|_{\square}$.

3.2. Some Facts About $\|.\|_{\square}$

(i)

$$\|F\|_{\square} := \sup_{f,g:[0,1]\to[0,1]} |\langle f, Fg\rangle|$$

(ii)

$$\sup_{f,g:[0,1]\to[0,1]} |\langle f, Fg\rangle| \geq \sup_{f:[0,1]\to[0,1]} |\langle f, Ff\rangle|.$$

(iii)

$$|\langle f, Fg\rangle| \geq \sup_{S\subseteq[0,1]} \left|\int_{S\times S} F(x,y)dxdy\right|.$$

(iv)

$$\sup_{S\subseteq[0,1]} \left|\int_{S\times S} F(x,y)dxdy\right| \geq \frac{1}{2}\|F\|_{\square}.$$

Proof of (i) *and* (iii):

(i)

$$\begin{aligned}
|\langle f, Fg\rangle| &= \int_0^1 f(x)F(g)dx \\
&= \int_0^1 f(x)\left(\int_0^1 F(x,y)g(y)dy\right)dx \\
&\leq \sup_{f,g:[0,1]\to[0,1]} \int_0^1\int_0^1 F(x,y)f(x)g(y)dydx
\end{aligned}$$

Therefore

$$\sup_{f,g:[0,1]\to[0,1]} \int_0^1\int_0^1 F(x,y)f(x)g(y)dydx = \sup_{S,T\subseteq[0,1]} \int_0^1 F(x,y)dydx = \|F\|_{\square}.$$

(iii)

$$\begin{aligned}
|\langle \chi_s, F(\chi_s)\rangle| &= \left|\int_0^1 \chi_s(x), F(\chi_s)(x)dx\right| \\
&= \int_0^1 \chi_s(x)\left[\int_0^1 F(x,y)(\chi_s(y)dy\right]dx \\
&= \left|\int_0^1\int_0^1 F(x,y)\chi_s(x)\chi_s(y)dydx\right| \\
&= \left|\int_{S\times S} F(x,y)\chi_{s\times s}(x,y)dxdy\right|.
\end{aligned}$$

This implies

$$\sup_{\chi_s:[0,1]\to[0,1]} |\langle \chi_s, F(\chi_s)\rangle| \geq \sup_{S\subseteq[0,1]} \left|\int_{S\times S} F(x,y)dxdy\right|$$

So

$$\sup_{f:[0,1]\to[0,1]} |\langle f, F(f)\rangle| \geq \sup_{S\subseteq[0,1]} \left|\int_{S\times S} F(x,y)dxdy\right|.$$

A version of the Szemeredi's Regularity Lemma can be formulated in this language as follows:

Lemma 3.1 (Weak Regularity Lemma, analytic form)**.** *For every* $W \in \mathcal{W}_0$ *and* $\varepsilon > 0$ *there is a step function* $W' \in \mathcal{W}_0$ *with at most* $\lceil 2^{2/\varepsilon^2} \rceil$ *steps such that* $\|W - W'\|_\square \leq \varepsilon$*.*

For every $W \in \mathcal{W}$ and every measurable partition $\mathcal{P}$ of $[0,1]$, let $W_{\mathcal{P}} : [0,1]^2 \longrightarrow \mathbb{R}$ denote the step function obtained by replacing its value at $(x,y) \in P_i \times P_j$ with the average of W over $P_i \times P_j$.

One can formulate the strong regularity lemma in this language as follows:

Lemma 3.2 (Strong Regularity Lemma, analytic form)**.** *For every* $\varepsilon > 0$ *there is an integer* k *such that for every function* $W \in \mathcal{W}_0$*, there is a partition* $\mathcal{P}$ *of* $[0,1]$ *into* k *sets of equal measure with the following property: For every set* $R \subseteq [0,1]^2$ *that is the union of at most* k^2 *rectangles, we have*

$$\left|\int_R (W - W_{\mathcal{P}})dxdy\right| \leq \varepsilon.$$

4. REGULARITY IN HILBERT SPACE

Lemma 4.1. *Let* $K_1, K_2, \ldots, K_n, \ldots,$ *be an arbitrary collection of nonempty subsets of a Hilbert Space* H. *For each* n, *let* $A(n) = \{\sum_{i=1}^{n} \gamma_i f_i : f_i \in K_i$ *and* $\gamma_i \in \mathbb{R}\}$. *Then, for every* $\varepsilon > 0$ *and* $f \in H$ *there is an* $m \leq \lceil 1/\varepsilon^2 \rceil$ *and* $\tilde{f} \in A(m)$ *such that for every* $g \in K_{m+1}$ *we have*

$$\left| \langle g, f - \tilde{f} \rangle \right| \leq \varepsilon \|f\| \|g\|.$$

The lemma is interesting in itself. As an application, consider the following case: Let $H = L^2([0,1])$. For each n let K_n be the set of all polynomials of degree at most 2^n. When we apply this lemma to this case, we get

Corollary 4.2. *Let* $\varepsilon > 0$ *and* $f \in L^2([0,1])$. *Then, there exists some polynomial* $p \in \mathbb{R}[x]$ *of degree* $d \leq 2^{\lceil 1/\varepsilon^2 \rceil}$ *such that*

$$\langle g, f - p \rangle \leq \varepsilon \|f\| \|g\|$$

for every polynomial g *of degree at most* $2d$.

Proof of the lemma. Let $\varepsilon > 0$ and $f \in H$. For each positive integer k, define

$$\alpha_k := \inf \left\{ \|f - \varphi\|^2 \mid \varphi \in A(k) \right\} = \inf_{\varphi \in A(k)} \|f - \varphi\|^2.$$

For each k, by the very definition of $A(k)$, zero function is in $A(k)$. Therefore, $\|f\| \in \{\|f - \varphi\|^2 \mid \varphi \in A(k)\}$. First, we observe that $A(1) \subseteq A(2) \subseteq A(3) \subseteq ... \subseteq A(k) \subseteq A(k+1)$. We then have

$$\|f\|^2 \geq \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_k \geq \alpha_{k+1} \ldots \geq 0.$$

We claim that there is some $m \leq \lceil 1/\varepsilon^2 \rceil$ such that

$$\alpha_m < \alpha_{m+1} + \varepsilon^2 \|f\|^2.$$

Observe that (α_n) is a decreasing sequence of real numbers bounded by 0 from below. Since $\alpha_1 \leq \|f\|^2$, for each $p \in \mathbb{N}$ we have $\alpha_1 - \alpha_{p+1} < \|f\|^2$.
Suppose that our claim is not true. In other words, let us assume that for each $m \leq \lceil 1/\varepsilon^2 \rceil$ we have $\alpha_m \geq \alpha_{m+1} + \varepsilon^2 \|f\|^2$. Now put $p = \lceil 1/\varepsilon^2 \rceil$. So

$$\begin{aligned} \alpha_1 &\geq \alpha_2 + \varepsilon^2 \|f\|^2 \\ \alpha_2 &\geq \alpha_3 + \varepsilon^2 \|f\|^2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \alpha_{p-1} &\geq \alpha_p + \varepsilon^2 \|f\|^2 \\ \alpha_p &\geq \alpha_{p+1} + \varepsilon^2 \|f\|^2 \end{aligned}$$

Adding these inequalities side by side and simplify (together with the inequality from the note above), we have $\|f\|^2 > \alpha_1 - \alpha_{p+1} \geq p \cdot \varepsilon^2 \|f\|^2$, which implies $p < \lceil 1/\varepsilon^2 \rceil$. This contradiction implies that our claim is true. So, there is some $m \leq \lceil 1/\varepsilon^2 \rceil$ such that $\alpha_m < \alpha_{m+1} + \varepsilon^2 \|f\|^2$. This means that (by the definition of α_m), there is some $\varphi \in A(m)$ such that

$$\|f - \varphi\|^2 \leq \alpha_{m+1} + \varepsilon^2 \|f\|^2.$$

Now consider any $g \in K_{m+1}$. By the definition of α_{m+1}, we have

$$\|f - (\varphi + tg)\|^2 \geq \alpha_{m+1} \geq \|f - \varphi\|^2 - \varepsilon^2 \|f\|^2$$

for every $t \in \mathbb{R}$. Hence we have

$$\|g\|^2 t^2 - 2\langle g, f-\varphi\rangle t + \varepsilon^2 \|f\|^2 \geq 0$$

for every real t. It follows that

$$4\langle g, f-\varphi\rangle^2 \leq 4\varepsilon^2 \|f\|^2 \|g\|^2$$

which implies that

$$\langle g, f-\varphi\rangle^2 \leq \varepsilon^2 \|f\|^2 \|g\|^2.$$

This is the required result.

4.1. RLHS Implies WRL

Let $H = L^2([0,1]^2)$ and let K_n denote the set of characteristic functions of measurable squares in $[0,1]^2$. If $f \in \mathcal{W}_0$, then $\varphi \in A(k)$ is a step function with at most 2^k steps. Hence if $W \in \mathcal{W}_0$, we get a step function W^* with at most $2^{\lceil 1/\varepsilon^2 \rceil}$ $(k = \lceil 1/\varepsilon^2 \rceil)$ steps such that for every measurable set $S \subseteq [0,1]$, by the previous lemma, we have $\forall g \in K_{m+1}, |\langle g, W - W^* \rangle| \leq \varepsilon \|W\| \cdot \|g\|$. Putting $g = \chi_{S\times S}$ implies that

$$\left| \int_{S\times S} (W - W^*) dxdy \right| \leq \varepsilon.$$

(Here, we consider the characteristic function of $S \times S$ as an element of K_{k+1} and use the definition of the inner product.)

From Lovazs-Szegedy (also on Section 3.2, fact (iv)), we know that

$$\sup_{S\subseteq[0,1]} \left| \int_{S\times S} W(x,y) dxdy \right| \geq \frac{1}{2} \|W\|_\square.$$

So

$$\|W - W^*\|_\square \leq 2 \sup_{S \subseteq [0,1]} \left| \int_{S \times S} (W - W^*) dxdy \right| \leq 2\varepsilon,$$

and this implies that for any two measurable subsets S, T of $[0,1]$, we have

$$\left| \int_{S \times T} (W - W^*) dxdy \right| \leq 2\varepsilon.$$

Clearly, this implies the Weak Regularity Lemma.

4.2. The Graph Theoretical Version of The WRL

Let $G = (V, E)$ be a graph on n vertices. Let $A = (a_{ij})$ be the corresponding adjacency matrix of G. For each entry a_{ij}, depending on whether or not there is an edge between the corresponding vertices, we have 0 or 1. Now replace each entry a_{ij} by a square of size $\frac{1}{n} \times \frac{1}{n}$ with the constant function a_{ij} on this square. This is a step function which will be represented as W_G. Let $\mathcal{A}$ be the algebra of subsets of $[0,1]$ generated by the intervals corresponding to the vertices of G. Let K_n be the set of characteristic functions of $S \times S$ where $S \in \mathcal{A}$. Analogously to the proof of the RLHS, we get a partition $\mathcal{P} = \{S_1, S_2, ..., S_n\}$ of $[0,1]$ into sets in $\mathcal{A}$ such that

$$\left| \int_{S \times T} (W_G(x,y) - (W_G)_{\mathcal{P}}(x,y)) dxdy \right| \leq 2\varepsilon$$

for all $S, T \in \mathcal{A}$. Since, we define the step function $(W_G)_{\mathcal{P}}$ on a pair $(S_i, S_j) \in \mathcal{P}$ as the average of W_G over (S_i, S_j), if we call the set of vertices corresponding to S_i as V_i, then the value of $(W_G)_{\mathcal{P}}$ on $S_i \times S_j$ is

$$\frac{\int_{S_i \times S_j} W_G(x,y)}{|S_i|\,|S_j|} = \frac{\dfrac{1}{n^2} e(V_i, V_j)}{\dfrac{1}{n^2}\,|V_i|\,|V_j|} = \frac{e(V_i, V_j)}{|V_i|\,|V_j|} = d(V_i, V_j).$$

Now, assume that analytic form of WRL is true. Let $W_G \in \mathcal{W}_0$ be the function which was defined as above. By Alon *et al.*, 2003, we can replace W' by $(W_G)_{\mathcal{P}}$ at the cost of increasing the error by a factor of at most 2. For a given $\varepsilon > 0$, for all $S, T \subseteq V$ let $S_a, T_a \in \mathcal{A}$ be the corresponding elements. Then,

$$\int_{S_a \times T_a} W_G(x,y)dxdy = \frac{1}{n}\frac{1}{n}e_G(S,T)$$

and

$$\begin{aligned}
\int_{S_a \times T_a} (W_G)_{\mathcal{P}}(x,y)dxdy &= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}\int_{(S_i \cap S_a)\times(S_j \cap T_a)} (W_G)_{\mathcal{P}}(x,y)dxdy\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\frac{1}{n^2}e(V_i,V_j)}{\frac{|V_i|}{n}\frac{|V_j|}{n}} |S_i \cap S_a| \, |S_j \cap T_a| \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} d(V_i,V_j)\frac{1}{n}|V_i \cap S| \frac{1}{n}|V_j \cap T| \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d(V_i,V_j)|V_i \cap S| \, |V_j \cap T| \\
&= \frac{1}{n^2}e_{\mathcal{P}}(S,T).
\end{aligned}$$

So,

$$\left|\int_{S_a \times T_a} (W_G - (W_G)_{\mathcal{P}})dxdy\right| = \frac{1}{n^2}|e(S,T) - e_{\mathcal{P}}(S,T)| \leq 2\varepsilon.$$

Therefore, $|e(S,T) - e_{\mathcal{P}}(S,T)| \leq 2\varepsilon n^2 \leq \varepsilon |V|^2$.

Note: It is reasonable to assume that $|V| \geq \sqrt{2} \cdot |\mathcal{P}|$. Otherwise many elements of the partition must be singletons.

4.3. SRL as a Result of RLHS

Let us define a sequence $(a_i)_{i \in \mathbb{N}^+}$ of positive integers as follows; $a_1 = 1$ and $a_{k+1} = 2^{a_1^4 \dots a_{k+1}^4}$ and then apply RLHS to the Hilbert Space $L_2([0,1]^2)$ and W be a function as before, but choose K_n to be the set of step functions with at most a_n steps. RLHS gives us a function W^*, which is a step function with at most $m = a_1 a_2 \dots a_k$ steps. Let $S_1, \dots S_m$ be these steps. Then, this step function has the property that for every step function U with at most a_{k+1} steps, we have

$$\left| \int_{[0,1]^2} U(W - W^*) dx dy \right| \leq \varepsilon.$$

Let us further partition each S_i into an appropriate number of sets of measure $\frac{1}{m^2}$ and so we will have a remainder with measure less than $\frac{1}{m^2}$. If we combine all the remainders into a single set, its measure will be less than $m\frac{1}{m^2} = \frac{1}{m}$. Again we partition this set into sets of size $\frac{1}{m^2}$. There will be at most m such sets. Call them recyled sets. In this way, we get a partition $\mathcal{P} = \{T_1, \dots, T_{m^2}\}$ of $[0,1]$ into m^2 parts of equal size. Note that exactly m parts of them are recyled ones. Now, let $R \subseteq [0,1]^2$ be a set that is the union of at most m^2 rectangles. From R, if we remove all the points in $T_i \times T_j$ where either T_i or T_j is a recycled set, the remaining set $R^{'}$ is again he union of at most m^2 rectangles and since the sum of the measure of such $T_i \times T_j$'s is $2m\frac{1}{m^2} = \frac{2}{m} < \varepsilon$, measure of $R \setminus R^{'}$ is less than ε.

Clearly, the characteristic function of $R^{'}$ is a step function with at most $2^{m^4} \leq s(k+1)$ steps, so by RLHS we have

$$\left| \int_{R^{'}} (W - W^*) dx dy \right| \leq \varepsilon.$$

Now, consider

$$\left|\int_{R'}(W_{\mathcal{P}} - W^*)dxdy\right|.$$

If none of T_i, T_j is a recyled set, then both W^* and $W_{\mathcal{P}}$ are constant on $T_i \times T_j$, so including or excluding $T_i \times T_j$ from $R^{'}$ will not decrease the value we are consider. Doing this for every pair of such T_i, T_j's, we obtain a set $R^{''}$ which is the union of some certain $T_i \times T_j$'s where none of T_i, T_j's are recycled. By the definiton of $W_{\mathcal{P}}$ and the result of RLHS we have

$$\left|\int_{R''}(W_{\mathcal{P}} - W^*)dxdy\right| = \left|\int_{R''}(W - W^*)dxdy\right| \leq \varepsilon$$

and so

$$\left|\int_{R'}(W_{\mathcal{P}} - W^*)dxdy\right| \leq \varepsilon.$$

Therefore

$$\left|\int_{R'}(W - W_{\mathcal{P}})dxdy\right| \leq \left|\int_{R'}(W - W^*)dxdy\right| + \left|\int_{R'}(W_{\mathcal{P}} - W^*)dxdy\right| \leq 2\varepsilon.$$

Hence

$$\left|\int_{R}(W - W_{\mathcal{P}})dxdy\right| \leq 3\varepsilon.$$

Given $\varepsilon > 0$ in the state of analytic form of SRL, take $\dfrac{\varepsilon}{3}$ in the state of RLHS. This concludes the proof.

5. REGULARITY LEMMA BY COVERING BALLS

In this chapter, by a distance function we mean a semi-metric.

First, we recall what a weak Szemeredi Partition is: A measurable partition P of $[0,1]$ is said to be a Weak Szemeredi Partition for W with error ε if

$$\sup_{S\subseteq[0,1]}\left|\int_{S\times S}(W-W_{\mathcal{P}})\right|\leq\varepsilon.$$

Weak Regularity Lemma (analytic form) says that every W has a Weak Szemeredi Partition with error ε in at most $\lfloor 2^{2/\varepsilon^2}\rfloor$ classes. If $W\in\mathcal{W}, x_1, x_2\in[0,1]$ define

$$d_W^1(x_1,x_2):=\| W(x_1,\cdot)-W(x_2,\cdot)\|_2=\left(\int_0^1 |W(x_1,y)-W(x_2,y)|^2\,dy\right)^{1/2}$$

and this implies d_W^1 is a distance function on $[0,1]$.

$d_W^1(x_1,x_2)=\| W(x_1,\cdot)-W(x_2,\cdot)\|_2$ is a semi-metric.

$W\in\mathcal{W}$ can be considered in several ways. As in chapter 3, W can be identified with an operator

$$W(f)(x)=\int_0^1 W(x,y)f(y)dy.$$

$T_W(:=W):L^2([0,1])\to L^2([0,1])$. We remark that $T_W\circ T_W$ is an operator of the form $T_{\tilde{W}}$. First let us find a $\tilde{W}$ such that $d_W=d_{\tilde{W}}^1$.

$$\begin{aligned}
(T_W \circ T_W(f))(x) &= \int_0^1 W(x,y) T_W(f)(y)dy \\
&= \int_0^1 W(x,y) \left(\int_0^1 W(z,y) f(y) dz \right) dy \\
&= \int_0^1 \left(\int_0^1 W(x,y) W(z,y) f(z) dy \right) dz \\
&= \int_0^1 \left(\int_0^1 W(x,y) W(z,y) dy \right) f(z) dz
\end{aligned}$$

Therefore we should consider

$$\tilde{W}(x,z) = \int_0^1 W(x,y) W(z,y) dy.$$

Now we have

$$d_W(x,x') = d^1_{\tilde{W}}(x,x') = \left(\int_0^1 (\tilde{W}(x,y) - \tilde{W}(x',y))^2 dy \right)^{1/2}$$

which is equal to

$$\left(\int_0^1 \left(\int_0^1 W(x,z) W(y,z) dz - \int_0^1 W(x',z) W(y,z) dz \right)^2 dy \right)^{1/2}.$$

Therefore $d^1_{\tilde{W}} = d_W$ is also a semi-metric on $[0,1]$.

Now we formulate the following;

Theorem 5.1. *Let $W \in \mathcal{W}_0$ and let $\mathcal{P}$ be a partition of $[0,1]$ into measurable sets.*

(i) *If $\mathcal{P}$ is a weak Szemeredi partition with error $\varepsilon^2/8$, then there is a set $S \subseteq [0,1]$ with $\mu(S) \leq \varepsilon$ such that, for each partition class, $P_i - S$ has diameter at most ε in the d_W metric.*

(ii) *If there is a set $S \subseteq [0,1]$ with $\mu(S) \leq (\varepsilon/5)^4$ such that, for each partition class, $P_i - S$ has diameter at most $(\varepsilon/5)^2$ in the d_W metric, then $\mathcal{P}$ is a weak Szemeredi partition with error ε.*

Proof. (i) Suppose that $\mathcal{P}$ is a Weak Szemeredi Partition with error $\varepsilon^2/8$. This means that $\|W - W_{\mathcal{P}}\|_{\square} \leq \varepsilon$. Let $R = W - W_{\mathcal{P}}$. If $x \in [0,1]$, define

$$F(x) = \int_0^1 \left(\int_0^1 R(x,s)W(s,z)ds \right)^2 dz.$$

Let us consider

$$\int_0^1 F(x)dx = \int_0^1 \int_0^1 \int_0^1 \int_0^1 R(x,t)R(x,s)W(s,z)W(t,z)dxdsdtdz.$$

If we fix z and t, since $-1 \leq R(x,t) \leq 1$ and $0 \leq W(s,z) \leq 1$, we have

$$\int_0^1 \int_0^1 R(x,t)R(x,s)W(s,z)dzds \leq \varepsilon^2/4.$$

This implies that $\int_0^1 F(x)dx \leq \varepsilon^2/4$. Hence, by the Tchebychev inequality, $\mu(\{x \in [0,1] : F(x) > \varepsilon/4\}) \leq \frac{4}{\varepsilon}\int_0^1 F(x)dx \leq \varepsilon$. Let $S = \{x \in [0,1] : F(x) > \varepsilon/4\}$. We then have $F(x) \leq \varepsilon/4$ for every $x \in [0,1] - S$. Therefore, if $x, y \in [0,1] - S$ are two points in the same partition class of $\mathcal{P}$, then $W_{\mathcal{P}}(x,s) = W_{\mathcal{P}}(y,s)$ for every $s \in [0,1]$, and hence

$$\begin{aligned}
d_W(x,y)^2 &= \int_0^1 \left(\int_0^1 (W(x,s) - W(y,s))W(s,z)ds \right)^2 dz \\
&= \int_0^1 \left(\int_0^1 (R(x,s) - R(y,s))W(s,z)ds \right)^2 dz \\
&= \int_0^1 \left(\int_0^1 R(x,s)W(s,z)ds - \int_0^1 R(y,s)W(s,z)ds \right)^2 dz \\
&\leq 2\int_0^1 \left(\int_0^1 R(x,s)W(s,z)ds \right)^2 dz + 2\int_0^1 \left(\int_0^1 R(y,s)W(s,z)ds \right)^2 dz \\
&= 2F(x) + 2F(y) \\
&\leq \varepsilon^2.
\end{aligned}$$

(ii) We want to show that $\|W - W_{\mathcal{P}}\|_{\square} \leq \varepsilon$. For this, it is enough to prove that for any $0-1$ valued function f, we have

$$\langle f, (W - W_{\mathcal{P}})f\rangle \leq \frac{1}{2}\varepsilon.$$

First, we write $f = f_{\mathcal{P}} + g$ where $f_{\mathcal{P}}(x)$ is obtained by replacing the average of f over the partition class containing x. Then, according to [Lovazs-Szegedy] which we have fact (ii) on Section 3.2 here, one has

$$\langle f, (W - W_{\mathcal{P}})f\rangle = \langle f + f_{\mathcal{P}}, Wg\rangle.$$

By the Cauchy-Schwarz inequality, we have

$$\langle f + f_{\mathcal{P}}, Wg\rangle \leq \|f + f_{\mathcal{P}}\|_2 \|Wg\|_2 \leq 2\|Wg\|_2.$$

Here, to say that $\|f + f_{\mathcal{P}}\|_2 \leq 2$, we use the definition of $f_{\mathcal{P}}$ and the fact that f is $0-1$ valued.

We need to understand $\|Wg\|$ better. First, we recall that W is symmetric. Hence we can write

$$\|Wg\|^2 = \langle g, W^2 g\rangle = \int_{[0,1]^3} g(x)W(x,y)W(y,z)g(z)dxdydz.$$

Let $\mathcal{P} = \{P_1, P_2, \ldots, P_k\}$ be our partition. Let $i \in [1,k]$. If $P_i \nsubseteq S$, let $x_i \in P_i - S$ be a fixed element. Otherwise define x_i to be 0. In this way, we get a function $\phi : [0,1] \longrightarrow [0,1]$ by

$$\phi(x) = x_i \text{ if } x \in P_i.$$

This function is clearly measurable. In fact, it is a simple function. If we put

$$W(x,y)W(y,z) = W(x,y)W(y,z) - W(x,y)W(y,\phi(z)) + W(x,y)W(y,\phi(z)),$$

we get that

$$\int_{[0,1]^3} g(x)W(x,y)W(y,z)g(z)dxdydz$$

is equal to the sum of

$$\int_{[0,1]^3} g(x)(W(x,y)W(y,z) - W(x,y)W(y,\phi(z)))g(z)dxdydz$$

and

$$\int_{[0,1]^3} g(x)W(x,y)W(y,\phi(z))g(z)dxdydz.$$

Let us first consider the second summand. By Fubini, we first integrate with respect to z. In this case, we can write

$$\int_{[0,1]^2} \left(\sum_{i=1}^{k} \int_{P_i} g(x)W(x,y)W(y,\phi(z))g(z)dz \right) dxdy.$$

Let us recall that ϕ is constant on each partition class. Hence the second summand is equal to

$$\int_{[0,1]^2} \left(\sum_{i=1}^{k} g(x)W(x,y)W(y,x_i) \int_{P_i} g(z)dz \right) dxdy.$$

But, by its definition $g = f - f_{\mathcal{P}}$ has integral zero over each P_i. Therefore the second summand is equal to 0. Moreover, by the Cauchy-Schwarz inequality, the first summand is smaller than

$$\left(\int_{[0,1]^3} g(x)^2 g(z)^2 dxdydz \right)^{\frac{1}{2}} \left(\int_{[0,1]^3} a(x,y,z)^2 dxdydz \right)^{\frac{1}{2}}$$

where $a(x,y,z) = W(x,y)W(y,z) - W(x,y)W(y,\phi(z))$ for all $x,y,z \in [0,1]$. Now we have

$$\int_{[0,1]^3} a(x,y,z)^2 dxdydz = \int_0^1 \left(\int_{[0,1]^2} a(x,y,z)^2 dxdy \right) dz$$

which is equal to

$$\int_0^1 d_W(z,\phi(z))^2 dz = \int_{[0,1]-S} d_W(z,\phi(z))^2 dz + \int_S d_W(z,\phi(z))^2 dz.$$

By hypothesis, if $z \in [0,1]-S$, we have $d_W(z,\phi(z)) \leq \left(\frac{\varepsilon}{5}\right)^2$ and if $z \in S$, we have $d_W(z,\phi(z)) \leq 1$. If we combine these with the fact that $\mu(S) \leq (\frac{\varepsilon}{5})^4$, we see that

$$\int_{[0,1]^3} a(x,y,z)^2 dxdydz \leq 2\left(\frac{\varepsilon}{5}\right)^4.$$

We also note that

$$\int_{[0,1]^3} g(x)^2 g(z)^2 dxdydz \leq 1.$$

Combining all of these we get

$$\int_{[0,1]^3} g(x)W(x,y)W(y,z)g(z)dxdydz \leq 2^{\frac{1}{2}}\left(\frac{\varepsilon}{5}\right)^2.$$

This means that

$$\|Wg\|_2 \leq \frac{1}{5}2^{\frac{1}{4}}\varepsilon \leq \frac{1}{4}\varepsilon.$$

We go back and get

$$\langle f + f_{\mathcal{P}}, Wg \rangle \leq 2\|Wg\|_2 \leq \frac{1}{2}\varepsilon.$$

Hence, we have

$$\langle f, (W - W_{\mathcal{P}})f \rangle \leq \frac{\varepsilon}{2}$$

and this finishes the proof.

$\square$

By the weak regularity lemma, we know that there always exist Weak Szemeredi Partitions If we apply this fact, we get the following

Corollary 5.2. *For every function* $W \in \mathcal{W}$ *and every* $\varepsilon > 0$*, there is a partition* $\mathcal{P} = \{P_0, P_1, \ldots, P_k\}$ *of* $[0,1]$ *into measurable sets with* $k \leq 2^{\lceil \frac{64}{\varepsilon^4} \rceil}$ *such that* $\mu(P_0) \leq \varepsilon$ *and, for* $1 \leq i \leq k$*,* P_i *has diameter at most* ε *in the* d_W *metric.*

6. CONCLUSION

We worked out the details of some parts of the [Lovasz and Szegedy]. So far, we have two proofs of Regularity Lemma in the analytic language. We also realized that some analytic versions of the Regularity Lemma implies some graph theoretical versions. Our motivation was and is to develop enough intuition to attack the same problems in the higher rank cases. What we mean by this can be explained as follows: A graph can contain two types of data, namely vertices and edges. Edges can be considered as a set of pairs of vertices. In a hypergraph, we have more general tuples. For example, a hypergraph of next level (say, 2) could contain vertices, edges, and triangles. Triangles could be represented by triples of vertices. This way, a hypergraph of rank 2 could be expressed as a stepfunction on $[0,1]^3$ instead of $[0,1]^2$. The cut norm can also be defined for this case. Just replace rectangles by generalized ones. In the analytic language, we believe some hypergraph version of the Regularity Lemma can be proved. But to do so, first an appropriate statement regarding to it should be produced and this needs further study. Furthermore, the relation between hypergraph version of a Regularity Lemma and the analytic form can be considered as a topic for further research.

APPENDIX:
SZEMERÉDI'S REGULARITY LEMMA:
A COMBINATORIAL PROOF

Theorem .1 (Szemeredi's Regularity Lemma)**.** *For every* $m \in \mathbb{N}^+$ *and* $\varepsilon > 0$*, there are integers* $M = M(\varepsilon, m)$ *and* K *such that every graph* (V, E) *of order greater than or equal to* K *has an* $\varepsilon -$ *regular partition* $P = (C, V_1, ...V_k)$ *with* $m \leq k \leq M$ *(Every graph has an* $\varepsilon -$*regular partition with a bounded number of classes). i.e This partition needs to satisfy the following properties;*

(i) $|C| \leq \varepsilon \cdot |V|$

(ii) $|V_1| = |V_2| = ... = |V_k|$.

(iii) $m \leq k \leq M$

(iv) *At most* $\varepsilon \cdot k^2$ *pairs* (V_i, V_j) *are not* $\varepsilon -$ *regular.*

Definition .2. *Let* $G = (V, E)$ *be a graph. We will define a function*

$$q' : \{(X, Y) \mid X \subseteq V, Y \subseteq V, X \cap Y = \varnothing\} \rightarrow [0, 1]$$

by

$$q'(X, Y) = \frac{|X| \cdot |Y|}{|V|^2} \cdot d(X, Y)^2 = \frac{e(X, Y)^2}{|V|^2 \cdot |X| \, |Y|}.$$

Note: Since $d(X, Y) \leq 1$, we have

$$q'(X, Y) \leq \frac{|X| \cdot |Y|}{|V|^2} \leq 1.$$

Definition .3. *For an equitable partition $P = (C, V_1, ...V_k)$ of the graph $G = (V, E)$, define*

$$q(P) := \sum_{1 \leq i < j \leq k} q'(V_i, V_j).$$

Note:

$$q(P) \leq \sum_{1 \leq i < j \leq k} \frac{|V_i| \cdot |V_j|}{|V|^2} \leq \frac{k(k-1)}{2} \frac{|V_1|^2}{|V|^2} \leqslant \frac{1}{2}$$

and this is an *upper bound* no matter what the partition size is.

Lemma .4. *Let X, Y be two disjoint subsets of V such that $X = \bigcup_i X_i$ and $Y = \bigcup_j Y_j$ where $X_i's$ and $Y_j's$ are both pairwise disjoint and the unions are finite. Then, we have*

$$q'(X, Y) \leq \sum_{i,j} q'(X_i, Y_j).$$

Proof. By Cauchy-Shwarz Inequality, we have

$$\left(\sum_k a_k \cdot b_k\right)^2 \leq \left(\sum_k a_k^2\right) \cdot \left(\sum_k b_k^2\right),$$

So

$$\sum_k b_k^2 \geq \frac{\left(\sum_k a_k \cdot b_k\right)^2}{\sum_k a_k^2}.$$

Let us assume $|P(X)| = s$ and $|P(Y)| = t$. Now, for each integer k with $1 \leq k \leq s \cdot t$, after ordering those pairs (X_i, Y_j), put $a_k = \sqrt{|X_i| \cdot |Y_j|}$ and $b_k = \sqrt{|X_i| \cdot |Y_j|} \cdot d(X_i, Y_j)$ into the previous inequality, we would have

$$\sum_{i,j} |X_i| \cdot |Y_j| \cdot d(X_i, Y_j)^2 \geq \frac{\left(\sum_{i,j} |X_i| \cdot |Y_j| \cdot d(X_i, Y_j)\right)^2}{\sum_{i,j} |X_i| \cdot |Y_j|}.$$

$$
\begin{aligned}
\sum_{i,j} q'(X_i, Y_j) &= \frac{1}{|V|^2} \cdot \left(\sum_{i,j} |X_i| \cdot |Y_j| \cdot d(X_i, Y_j)^2 \right) \\
&\geq \frac{1}{|V|^2} \frac{\left(\sum_{i,j} |X_i| \cdot |Y_j| \cdot d(X_i, Y_j) \right)^2}{\sum_{i,j} |X_i| \cdot |Y_j|} \\
&= \frac{1}{|V|^2} \frac{\left(\sum_{i,j} e(X_i, Y_j) \right)^2}{\sum_{i,j} |X_i| \cdot |Y_j|} \\
&= \frac{1}{(|V|)^2} \cdot \frac{e(X,Y)^2}{|X| \cdot |Y|} = q'(X,Y).
\end{aligned}
$$

□

Quality Partition Refinement

Suppose we have an equipartition $P_r = \{C_r, V_1, ..., V_{n_r}\}$ that is not $\varepsilon - regular$.(ie. For each $1 \leq i \leq n_r$, $|V_i| = s_r$ and assume $|C_0| < \frac{\varepsilon}{2} |V|$.) For a fixed i where $1 \leq i \leq n_r$, consider (V_i, V_j) where $1 \leq j \leq n_r$ but $j \neq i$. For an $\varepsilon - regular$ pair, put $W_{ij} = \varnothing$. If (V_i, V_j) is not $\varepsilon - regular$, then let W_{ij} be a subset of V_i that violates the definition of the $\varepsilon - regularity$ of this pair.(In this situation, of course W_{ij} will be a proper subset of V_i. For a fixed i, the number of all $(V_i, V_j)'s$, above is $n_r - 1$. For this V_i, consider the set of all those W_{ij}'s that are non-empty. Let

$$
u_{ij} = \bigcap_{j=1, j \neq i}^{n_r} D_{ij}
$$

where $D_{ij} = W_{ij}$ or $D_{ij} = V_i \backslash W_{ij}$. Now let U_i be the set of all those u_{ij}'s. The total number of elements of U is at most $2^{n_r - 1}$, each element in this set is a subset of V_i and u_{ij}'s are pairwise disjoint. In this way, we form a partition of V_i. Suppose we do this for all other V_i's in the same way and form their partitions. When we collect all of them, we come up a refinement P_r' of P_r. It is clear that $|P'| \leq n_r \cdot 2^{n_r - 1}$. At this point, it is possible that not all the parts of P'_r are of equal size. So we will pick some size s_{r+1} which we will be determined later. Cut through each element of P'_r into pieces of size s_{r+1} and send the left over elements to C_r to form C_{r+1}. Renaming or reenumerating pieces of size s_{r+1}, we end up our nextcoming partition $P_{r+1} = \left\{C_{r+1}, V_1, V_2, ..., V_{n_{r+1}}\right\}$.

Lemma .5. *Let V_x, V_y be two disjoint pieces in P_r that are not $\varepsilon - regular$. Moreover, let $X_1, ..., X_g$ and $Y_1, ..., Y_h$ are all the new pieces in P_{r+1} so that $X = \bigcup_{i=1}^{g} X_i \subseteq V_x$ and $Y = \bigcup_{i=1}^{h} Y_i \subseteq V_y$. Then*

$$\sum_{\substack{1\le i\le g\\ 1\le j\le h}} q'(X_i, Y_j) \ge q'(X, Y) + \frac{\varepsilon^4}{2}\frac{s_r^2}{|V|^2}.$$

Proof. Since V_x, V_y are not $\varepsilon - regular$, there exists $W_x \subset V_x, W_y \subset V_y$ with $|W_x| > \varepsilon |V_x|, |W_y| > \varepsilon |V_y|$ such that $|d(W_x, W_y) - d(V_x, V_y)| > \varepsilon$. (Note that W_x, W_y are among the $W_{i,j}$'s when we are constructing P'_r). Now, let w_x, w_y be the number of new pieces from P_{r+1} that fall within W_x, W_y respectively. By the very construction of these sets, we have;

(i) The number of vertices thrown from V_x, V_y, W_x and W_y to C_r are equal to $|V_x - X|, |V_Y - Y|, |W_x - s_{r+1}w_x|$ and $|W_y - s_{r+1}w_y|$ respectively.
(ii) Clearly, $|W_x - s_{r+1}w_x| \le |V_x - X|$ and $|W_y - s_{r+1}w_y| \le |V_y - Y|$.

Now, for a finite sequence of real numbers $x_1, ...x_n$, put $\mu = \dfrac{x_1 + x_2 + ... + x_n}{n}$. Then we have

$$\sum_{k=1}^{n} x_k^2 = \sum_{k=1}^{n} (x_k - \mu)^2 + n\cdot \mu^2.$$

(This is easy to verify, so we are skipping the proof of it) In this expression, if we put $x_k = e(X_i, Y_j)$, we will have

$$\sum_{\substack{1\le i\le g\\ 1\le j\le h}} e(X_i, Y_j)^2 = \sum_{\substack{1\le i\le g\\ 1\le j\le h}} \left(e(X_i, Y_j) - \frac{e(X, Y)}{g\cdot h}\right)^2 + g\cdot h\cdot \left(\frac{e(X, Y)}{g\cdot h}\right)^2.$$

This implies

$$\begin{aligned}\sum_{\substack{1\le i\le g\\1\le j\le h}} q'(X_i,Y_j) &= \frac{1}{s_{r+1}^2\cdot|V|^2}\cdot\sum_{\substack{1\le i\le g\\1\le j\le h}} e(X_i,Y_j)^2\\ &= \frac{1}{s_{r+1}^2\cdot|V|^2}\cdot\sum_{\substack{1\le i\le g\\1\le j\le h}}\left(e(X_i,Y_j)-\frac{e(X,Y)}{g\cdot h}\right)^2+\frac{e(X,Y)^2}{ghs_{r+1}^2\,|V|^2}.\end{aligned}$$

Note: Since, by our definition

$$q'(X,Y)=\frac{e(X,Y)^2}{g\cdot h\cdot s_{r+1}^2\,|V|^2},$$

the last term in the previous sum is nothing but $q'(X,Y)$.

By Quality Partition Refinement, we have

$$\sum_{\substack{1\le i\le g\\1\le j\le h}}\left(e(X_i,Y_j)-\frac{e(X,Y)}{g\cdot h}\right)^2\ge\sum_{\substack{X_i\in W_x\\Y_j\in W_y}}\left(e(X_i,Y_j)-\frac{e(X,Y)}{g\cdot h}\right)^2,$$

and if we put $x_k=e(X_i,Y_j)-\dfrac{e(X,Y)}{g\cdot h}$, $b_k=1$, then by Cauchy-Shwarz Inequality we have

$$\sum_{\substack{X_i\in W_x\\Y_j\in W_y}}\left(e(X_i,Y_j)-\frac{e(X,Y)}{g\cdot h}\right)^2\ge\frac{\left(\sum_{\substack{X_i\in W_x\\Y_j\in W_y}}\left(e(X_i,Y_j)-\frac{e(X,Y)}{g\cdot h}\right)\right)^2}{\sum_{\substack{X_i\in W_x\\Y_j\in W_y}}1^2}.$$

Now using basic counting principles, we have

$$\begin{aligned}
\frac{\left(\sum_{\substack{X_i\in W_x\\ Y_j\in W_y}}\left(e(X_i,Y_j)-\dfrac{e(X,Y)}{g\cdot h}\right)\right)^2}{\sum_{\substack{X_i\in W_x\\ Y_j\in W_y}}1^2} &= \left(\frac{1}{w_xw_y}\right)\left(e(W_x,W_y)-\frac{w_xw_ye(X,Y)}{gh}\right)^2\\
&= w_xw_y\left(\frac{(e(W_x,W_y)}{w_xw_y}-\frac{e(X,Y)}{gh}\right)^2\\
&= w_xw_ys_{r+1}^4\left(\frac{(e(W_x,W_y)}{s_{r+1}^2w_xw_y}-\frac{e(X,Y)}{s_{r+1}^2gh}\right)^2\\
&\geq \frac{1}{2}\left|W_x\right|\left|W_y\right|s_{r+1}^2\left(\frac{(e(W_x,W_y)}{\left|W_x\right|\left|W_Y\right|}-\frac{e(X,Y)}{\left|X\right|\left|Y\right|}\right)^2\\
&\geq \frac{1}{2}\left|W_x\right|\left|W_Y\right|s_{r+1}^2\varepsilon^2\\
&\geq \frac{1}{2}\varepsilon\left|V_x\right|\cdot\varepsilon\left|V_y\right|s_{r+1}^2\varepsilon^2\\
&= \frac{1}{2}\varepsilon^4s_r^2s_{r+1}^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{\substack{1\leq i\leq g\\ 1\leq j\leq h}}q'(X_i,Y_j) &\geq \frac{1}{s_{r+1}^2\left|V\right|^2}\cdot\frac{1}{2}\varepsilon^4s_r^2s_{r+1}^2+\frac{e(X,Y)^2}{ghs_{r+1}^2\left|V\right|^2}\\
&= \frac{1}{2}\frac{\varepsilon^4s_r^2}{\left|V\right|^2}+\frac{e(X,Y)^2}{\left|V\right|^2\left|X\right|\left|Y\right|}\\
&= \frac{1}{2}\frac{\varepsilon^4s_r^2}{\left|V\right|^2}+\frac{d(X,Y)s_r^2}{\left|V\right|^2}\\
&\geq \frac{\varepsilon^4}{2}\frac{s_r^2}{\left|V\right|^2}+q'(X,Y).
\end{aligned}$$

□

Proof. (Szemeredi's Regularity Lemma) If P_r needs to be refined, a nextcoming refined partition will be P_{r+1} and $\left|C_{r+1}\right|-\left|C_r\right|\leq s_{r+1}\left|P'(r)\right|\leq s_{r+1}n_r2^{n_r-1}$, pick $s_{r+1}=\frac{s_r}{4^{n_r}}$. So $\left|C_{r+1}\right|-\left|C_{r+}\right|\leq\frac{s_rn_r}{2^{n_r+1}}\leq\frac{|V|}{2^{n_0+1}}$. Therefore $\left|C_{r+1}\right|-\left|C_0\right|\leq\left|V\right|\cdot\left(\frac{1}{2^{n_0+1}}+\frac{1}{2^{n_1+1}}+...+\frac{1}{2^{n_0+1}}\right)\leq\frac{1}{2^{n_0}}\left|V\right|$. For a given $\varepsilon>0$, if we choose n_0 such that $\left|C_0\right|<\frac{\varepsilon}{2}\left|V\right|$ and $\frac{1}{2^{n_0}}<\frac{\varepsilon}{2}$, then $\left|C_{r+1}\right|\leq\left|C_0\right|+\frac{|V|}{2^{n_0}}<\frac{\varepsilon}{2}\left|V\right|+\frac{\varepsilon}{2}\left|V\right|=\varepsilon\left|V\right|$. (Here the value n_0 gives us a lower bound m for the partition size.)

For a not $\varepsilon - regular$ pair (X, Y) in a partition P_r of V, we know from the lemma A.5 that

$$\sum_{\substack{1 \le i \le g \\ 1 \le j \le h}} q'(X_i, Y_j) \ge \frac{\varepsilon^4}{2} \frac{s_r^2}{|V|^2} + q'(X, Y).$$

If the last condition of the lemma does not hold, in other words, there are at least εn_r^2 pairs in the partition P_r, we make a quality refinement partition P_{r+1}. When we compare $q(P_{r+1})$ and $q(P_r)$, for each not $\varepsilon - regular$ pair $\frac{\varepsilon^4}{2}$ needs to be added. So, $q(P_{r+1}) \ge q(P_r) + \varepsilon n_r^2 \cdot \frac{\varepsilon^4}{2} = q(P_r) + \frac{\varepsilon^5}{2}$. We know that q is bounded from above by $\frac{1}{2}$. Since $q(P_{r+1}) - q(P_r) \ge \frac{\varepsilon^5}{2}$, there can not be more than

$$\frac{\frac{1}{2}}{\frac{\varepsilon^5}{2}} = \frac{1}{\varepsilon^5}$$

quality partition refinements of V. This means that the total number of quality partition refinements is less than or equal to $\lfloor 1/\varepsilon^5 \rfloor$.

What can we say about an upper bound M?

Since $n_r s_r \ge n_{r+1} s_{r+1}$, this implies that $n_{r+1} \le \frac{n_r s_r}{s_{r+1}}$. By our choice of $s_{r+1} = \frac{s_r}{4^{n_r}}$, we have $n_{r+1} \le n_r 4^{n_r}$. Let us consider a function f defined on the set of natural numbers, such that for each $x \in \mathbb{N}$, $f(x) = x4^x$. For our problem, let $\theta = \lfloor 1/\varepsilon^5 \rfloor$. Obviously, the number of quality partition refinements needed is $\le \theta$. Since $n_1 \le f(n_0), n_2 \le f(n_1), ..., n_\theta \le f(n_{\theta-1})$, we have $n_\theta \le f(f(n_{\theta-2})) \le ... \le \underbrace{(f \circ f \circ ... \circ f)}_{composition\ \theta\ times}(n_0)$. Hence, clearly $M = \underbrace{(f \circ f \circ ... \circ f)}_{composition\ \theta\ times}(n_0)$ works as an upper bound for the size of the partitions.

Note. We have the following remarks:

(i) This value M is of tower type.

(ii) In most of the cases, it may be possible to find smaller upper bounds.

(iii) The upper bound M depends on $m = n_0$ and ε. So $K \geq M$ will be fine. $\square$

REFERENCES

Lovász, L. and B. Szegedy, 2007, Szemerédi's Lemma for the Analyst, *Geometric and Functional Analysis*, Vol. 17, pp. 252-270.

Alon, N., W. Fernandez de la Vega, R. Kannan and M. Karpinski, 2003, Random Sampling and Approximation of MAX-CSPs, *Journal of Computer and System Sciences*, Vol. 67, pp. 212-243.

Gowers, W. T., 2007, Hypergraph Regularity and the Multidimensional Szemerédi Theorem, *Annals of Mathematics*, Vol. 166, pp. 897-946.

Tao, T., 2006, A Variant of the Hypergraph Removal Lemma, *Journal of Combinatorial Theory*, Series A, Vol. 113, pp. 1257-1280.

REFERENCES NOT CITED

Tao, T., 2006, Szemerédi's Regularity Lemma Revisited, *Contributions to Discrete Mathematics*, Vol. 1, pp. 8-28.

Bollobás, B., 1998, *Modern Graph Theory*, Springer GTM 184, New York.

Diestel, R., 2000, *Graph Theory*, Springer GTM 173, New York.

Royden, H. L., 1968, *Real Analysis*, Macmillan Publishing Co., New York.

Marcus, A., 2008, The Szemerédi Regularity Lemma, http://cs-www.cs.yale.edu/homes/marcus/papers/srl.pdf.